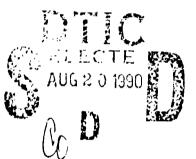AD-A226 006

Investigations of Naturalistic Decision Making and the Recognition-Primed Decision Model

Gary A. Klein and Roberta Calderwood

Klein Associates, Inc.

for

Contracting Officer's Representative
Judith Orasanu

DTIC
ELECTE
AUG 2 0 1990
D

Basic Research
Michael Kaplan, Director

July 1990

United States Army
Research Institute for the Behavioral and Social Sciences

U.S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency Under the Jurisdiction
of the Deputy Chief of Staff for Personnel
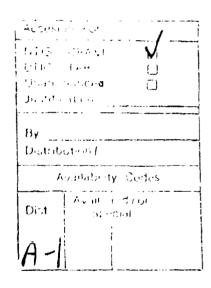
EDGAR M. JOHNSON
Technical Director

JON W. BLADES
COL, IN
Commanding

Research accomplished under contract for
the Department of the Army

Klein Associates, Inc.

Technical review by

George Lawton

NOTICES

REPORT DOCUMENTATION PAGE

. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
classified	--

. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT
	Approved for public release;
. DECLASSIFICATION / DOWNGRADING SCHEDULE	distribution is unlimited.

PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
-TR-88-02Z	ARI Research Note 90-59

. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
ein Associates, Inc.	--	U.S. Army Research Institute for the Behavioral and Social Sciences

ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code)
0 Livermore Street	5001 Eisenhower Avenue
llow Springs, OH 45387	Alexandria, VA 22333-5600

NAME OF FUNDING / SPONSORING ORGANIZATION U.S. Army Research stitute for the Behavioral d Social Sciences	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
	--	MDA903-85-C-0327

ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS			
01 Eisenhower Avenue exandria, VA 22333-5600	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
	6110CB	74F		

TITLE (Include Security Classification)
vestigations of Naturalistic Decision Making and the Recognition-Primed cision Model

PERSONAL AUTHOR(S)
ein, Gary A.; and Calderwood, Roberta.

. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
nal	FROM 85/07 TO 88/C7	1990, July	61

SUPPLEMENTARY NOTATION
ntracting Officer's Representative, Judith Orasanu

COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)	
FIELD	GROUP	SUB-GROUP	Decision making	Natural reasoning expertise
			Decision strategies	Time-pressure
			Predecision processes	Think-aloud protocols

ABSTRACT (Continue on reverse if necessary and identify by block number)
) This monograph reviews 3 years of research that explores how experienced personnel make cisions in operational settings characterized by real-time information processing, shifting als, and high-risk consequences.

The study combined field studies with experiments designed to test specific hypotheses. udy domains were selected so that findings would have high potential for generalizing to litary command-and-control decision making. Researchers carried out critical decision terviews with experienced personnel, including urban fire ground commanders, wildland fire cident commanders, and U.S. Army tank platoon leaders. Interviews were designed to elicit formation about the cues, goals, and option evaluation strategies used by these personnel. sed on these interviews, the relationships among such factors as time pressure, experience vel, and group interactions were explored.

(Continued)

DISTRIBUTION / AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION	
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	Unclassified	
. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
Michael Drillings	(202) 274-8722	PERI-BR

Form 1473, JUN 86   Previous editions are obsolete.   SECURITY CLASSIFICATION OF THIS PAGE

UNCLASSIFIED

ARI Research Note 90-59

19. ABSTRACT (Continued)

The results of these studies have been used to develop a Recognition-Primed Decision (RPD) model of decision making. This model contrasts with current normative and prescriptive models of decision making, and the implications of this alternative framework are explored. The findings that we consider most important are

- Experienced decision makers come to rely more on situation assessment, while novices rely more on option evaluation strategies.

- Situation assessment seems to involve schematic or prototypical knowledge of cues, goals, and expectancies that apply to a given class of events. Current cognitive research paradigms have not addressed how complex decision events are classified.

- Whereas experts and novices notice the same cues in a situation, novices draw fewer inferences based on these cues. Novices tend to miss the tactical implications of situational cues.

- At least in the domains studied here, decisions are most likely to be made without any conscious deliberation between option alternatives.

- When deliberation does occur, decision makers are more likely to use serial evaluation strategies than concurrent evaluation of options. Serial strategies appear to offer a means of minimizing the calculational burden, as well as maximizing the speed with which a decision may be implemented.

- Serial evaluation is associated with satisfying rather than optimizing strategies, and is preferred under time-pressured conditions.

- Options are frequently evaluated through the use of images or a "mental model" that operates as a simulation for judging whether an option will be successful in a specific case.

- Expert decision makers rely on a process of "progressive deepening" or reasoning into the future.

- Analogical reasoning is infrequently reported, which suggests that the processes involved in selecting and using analogues are relatively automatic and unconscious.

- When analogues are used (often to address aspects of a problem that are not routine), they are critical to option selection. Thus, inappropriate analogues are a primary cause of errors.

- Time pressure does not affect the quality of decisions made by experts as much as it affects novices because experts rely more heavily on rapid recognitional processes.

INVESTIGATIONS OF NATURALISTIC DECISION MAKING AND THE RECOGNITION-PRIMED DECISION MODEL

EXECUTIVE SUMMARY

This monograph reviews 3 years of research concerned with how experienced personnel make decisions in operational settings characterized by real-time information processing, shifting goals, and high-risk consequences.

The study combined field studies with experiments designed to test specific hypotheses. Study domains were selected so that findings would have high potential for generalizing to military command-and-control decision making. Critical Decision interviews were carried out with experienced personnel including urban fire ground commanders, wildland fire incident commanders, and U.S. Army tank platoon leaders. Interviews were designed to elicit information about the cues, goals, and option evaluation strategies used by these personnel. Based on these interviews, the relationships among such factors as time pressure, experience level, and group interactions were explored.

The results of these studies have been used to develop a Recognition-Primed Decision (RPD) model of decision making. This model contrasts with current normative and prescriptive models of decision making, and the implications of this alternative framework are explored. The findings that we consider most important are

- Experienced decision makers come to rely more on situation assessment, while novices rely more on option evaluation strategies.

- Situation assessment seems to involve schematic or prototypical knowledge of cues, goals, and expectancies that apply to a given class of events. Current cognitive research paradigms have not addressed how complex decision events are classified.

- Whereas experts and novices notice the same cues in a situation, novices draw fewer inferences based on these cues. Novices tend to miss the tactical implications of situational cues.

- At least in the domains studied here, decisions are most likely to be made without any conscious deliberation between option alternatives.

- When deliberation does occur, decision makers are more likely to use serial evaluation strategies than concurrent evaluation of options. Serial strategies appear to offer a means of minimizing the calculational burden, as well as maximizing the speed with which a decision may be implemented.

- Serial evaluation is associated with satisfying rather than optimizing strategies, and is preferred under time-pressured conditions.

- Options are frequently evaluated through the use of images or a "mental model" that operates as a simulation for judging whether an option will be successful in a specific case.

- Expert decision makers rely on a process of "progressive deepening" or reasoning into the future.

- Analogical reasons is infrequently reported, which suggests that the processes involved in selecting and using analogues are relatively automatic and unconscious.

- When analogues are used (often to address aspects of a problem that are not routine), they are critical to option selection. Thus, inappropriate analogues are a primary cause of errors.

- Time pressure does not affect the quality of decisions made by experts as much as novices, because experts rely more heavily on rapid recognitional processes.

INVESTIGATIONS OF NATURALISTIC DECISION MAKING AND THE RECOGNITION-PRIMED
DECISION MODEL

CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

.

INVESTIGATIONS OF NATURALISTIC DECISION MAKING AND THE
RECOGNITION-PRIMED DECISION MODEL

I. INTRODUCTION

This is the final report[1] for the program of research entitled

"Analogical Decision Making" sponsored by the U. S. Army Research Institute

for the Behavioral and Social Sciences which began in July, 1985 and ended in

July, 1988 (MDA903-85-C-0327). The goal of the project was to investigate

decision making in operational settings and to develop theory relevant to Army

needs. The research approach has relied primarily on obtaining and analyzing

verbal protocols of decision events from experienced personnel. We have

refined and modified both our data gathering and our analytic methods

throughout the course of the project, so that the methodological development

has formed a significant part of the research effort. Throughout the project

we have attempted to blend features of naturalistic field investigation with

the theory and hypothesis testing of behavioral decision making and cognitive

psychological research.

Seven separate projects were carried out under this program. Field

studies of command-and-control decision making included one on fire ground

command decision making, one on wildland fire incident command, and one on

U.S. Army tank platoon battle management. An experiment designed to test the

effect of time pressure on decision quality was carried out using chess as the

decision task. One project reviewed protocols from all of the field studies

for evidence of analogical reasoning, and proposed a preliminary taxonomy of

the functions of analogues in decision making. One project involved an

extensive literature review relevant to the categorization of natural decision

events. The review formed the basis for a doctoral dissertation proposed to

---

research the effect of context on event classification. The most recent study obtained think-aloud protocols of fire ground command decisions during simulated incidents. This effort allowed experimental verification of hypotheses generated on the basis of previously obtained interview data.

All but one of these projects have been reported previously in technical reports and published articles. Rather than providing an extensive review of each of these projects, it seems more appropriate to provide an overview of those findings that have had the most impact on theory development or applications. Complete citations for each of the referenced studies are provided, but the interested reader may also wish to refer to the titles and summaries of these reports that are provided in Section IV. The single previously unpublished study is included as Appendix I.

The remainder of this report is organized by topics and will not attempt to explain experimental procedures and results in detail or retain the chronological development of the ideas.

The Need for Descriptive Models

The idea that provided the major impetus for the research in this project was that models and research methods in behavioral decision research have been too focused on the analytic processes involved in comparing and selecting from among a predefined set of options. Many laboratory decision tasks are based on some variation of a gamble between two clearly defined bets or alternatives. Even in tasks where multiple options are available and multiple attributes are considered in the evaluation, both the options and the evaluation dimensions are well-defined and remain constant over the course of the decision problem.

2

Clearly, "real life" decisions rarely come so neatly packaged. Before alternatives can be evaluated or even identified, the decision must be framed or structured in some way that enables relevant goals to be identified and appropriate options to be generated. Moreover, the structure of the decision may change over time as events change and/or a new understanding of a decision problem is achieved. Although the limitations of standard decision research paradigms have long been recognized, there remains a lack of systematic research relevant to dynamic and ill-defined tasks (e.g. Elehmer & Allard, 1986; Edwards, 1962; Gettys, 1983; Rapoport & Wallsten, 1972).

Our primary interest has been on task domains that share the essential characteristics of command-and-control decision making: involving high-level integration of near real-time information for the purpose of deciding how best to utilize force application in a "battle" environment under varying degrees of uncertainty and time pressure (cf. Wohl, Entin, Kleinman, & Pattipati, 1984). Consider these examples of decision events:

> An incident commander is charged with controlling a fire raging through 18,000 acres of forest and range land. This fire has been burning for several days when a fortunate break in weather conditions allows a chance to renew attack efforts which have had to be largely defensive for the past two days. In consultation with his five experienced staff officers, he begins a session that will determine where to place a fire control line. Reports from air and ground surveillance are used as a basis for considering various placements of the direct attack, weighing such factors as the line length required, equipment and manpower proximity, whether or not bulldozers could be used or hand crews would be required, danger if efforts fail, the political ramifications of letting the fire burn a particular area, the effects of renewed bad weather.... A decision must be reached in this 3 a.m. session before the day's front-line attack crews are given orders at 6 a.m.

> An urban fire ground commander is called to the scene of a two-story wood frame building fire in a residential area. Reports from citizens at the scene indicate that all of

the residents are out of the building so he is able to
focus attention on setting fire control strategy. The
fire is already well-involved, and judging from the
location of the visible flames, color of the smoke, and
smoke concentrations, looks to have started in a
downstairs room, spreading very rapidly into one of the
upstairs bedrooms with very little horizontal spread. The
rate of spread could indicate a combustible fuel source
making the inside attack more dangerous and the prognosis
poorer for saving the structure. Standard operating
procedure is to order roof ventilation and an outside
attack until exposures are protected. However, the
vertical spread is so rapid that it looks as if the roof
may self-ventilate and an aggressive inside attack may
save the bulk of the structure and contents. He hates to
see these people lose everything, but he's seen these
things go real sour.... He must decide in less than one
minute whether to order all crews inside on hose lines or
wait for the truck crew to ventilate.

These decision makers must work quickly to clarify the nature of the
situation based on their own experience and training and they must decide
which of several conflicting goals should be given priority. The situation
may change as a result of their own actions or other events, so events must be
constantly monitored and reassessed.

Many of the issues of decision research paradigms are simply not directly
relevant to these types of decisions. We have found that the very language of
decision models is difficult to translate into operational settings. In one
of our earlier studies of urban fire ground commanders (FGCs) (Klein,
Calderwood, & Clinton-Ci.... , 1985) we were surprised to find the commanders
rejecting the notions that they were "making choices," "considering
alternatives," or "assessing probabilities." They saw themselves as acting
and reacting on the basis of prior experience, and generating, monitoring, and
modifying plans to meet the needs of the situations. Because we found no
evidence for extensive option generation there was little chance to observe
tradeoffs between the utilities of outcomes. Nor could we see any way to

4

apply the concept of optimal choice. It appeared that a search for an optimal choice could stall the FGCs long enough to lose control of the firefighting operations. The FGCs were more concerned with identifying actions that were "workable," "timely," and "cost-effective."

We originally proposed that understanding of these types of decision events could be increased by focusing on the natural reasoning strategies being used. An early hypothesis was that analogical reasoning was a primary basis for making decisions, based on previous work in inference and predictions (Klein & Weitzenfeld, 1982; Weitzenfeld, 1984). In making predictions, an individual frequently establishes a comparison case based on the similarity of the case to a target case. Similarity is not based on featural matching (e.g. Tversky, 1977), but on an overall judgment about whether the comparison cases contains the relevant causal factors. Adjustments are then made on the basis of differences between the target and comparison cases in order to make a prediction or inference.

Although we later determined that analogical reasoning was too narrow to account for the decision processes we were describing, the theoretical framework that we have developed retains the idea that decision making starts with an understanding of a situation based on previous experiences and knowledge.

We further proposed that laboratory methods using simplified tasks and inexperienced decision makers were an inherently inadequate basis for building models that would have applications to natural decision tasks. We wished to find methods that allowed the contextual constraints that are normally operating to be apparent. This seems to us the best way of generating potentially important hypotheses and for increasing the fit between theory and

5

practice. We settled on a quasi-naturalistic approach that has generated a rich source of data for generating hypotheses and suggesting fruitful avenues of research.

The next sections describe the theoretical framework, the methods that were developed within this program, and some of the major implications of the model for future research and applications.

II. RECOGNITION-PRIMED DECISIONS

Although the FGCs we studied denied making decisions in the traditional sense of "selecting an option," they were clearly making choices and judgments that affected the course of events. However the FGCs insisted that they rarely deliberated about the advantages and disadvantages of the different options. Instead the FGCs relied on their abilities to recognize and appropriately classify a situation. Once they knew it was "that" type of case, they usually also knew the typical way of reacting to it. They would use the available time to evaluate an option's feasibility before implementing it. Imagery might be used to "watch" the option being implemented, to discover if anything important might go wrong. If problems were foreseen then the option might be modified or rejected altogether.

For this task environment, this recognitional strategy appears to be very efficient. The proficient FGCs we studied could use their experience to generate a workable option as the <u>first</u> to consider. If they had tried to generate a large set of options and then systematically evaluate them, it is likely that the fires would have gotten out of control before they could make any decisions.

The Recognition-Primed Decision (RPD) model is presented in Figure 1. It shows the proficient decision maker becoming aware of events that have

Figure 1: Recognition-Primed Decision (RPD) Model

occurred, and relying on experience to recognize these events as largely typical. The simplest case is one in which the situation is recognized and the standard reaction is implemented. A somewhat more complex case is one in which the decision maker performs some conscious evaluation of the reaction, often using imagery to uncover problems. The most complex case is one in which the evaluation reveals flaws requiring modifications, or the option is judged inadequate and is rejected in favor of another typical reaction.

The evaluation function can operate at several levels. The verification/nonverification of expectancies serves to alert the decision maker that the situational understanding is wrong, and that it is time to rethink it and gather more information. In addition, individual actions are evaluated through progressive deepening to see how they will turn out.

This model clearly includes aspects of problem solving and judgment along with decision making. In naturalistic settings it is rare to find one without the others (e.g., see Wohl, 1981). There are three features of the model that will be discussed: (a) situation assessment, (b) serial evaluation, and (c) progressive deepening.

Situation Assessment

The experts we studied are able to quickly determine if a case poses any new challenges. Their experience with a wide variety of cases assures that most problems they encounter will have many features similar to what they have seen before. In other words, they form a situation assessment based on judgments of prototypicality (Rosch & Mervis, 1975) that activates appropriate "schemas" or "scripts" (Bartlett, 1932; Kintsch & van Dijk, 1975). If the judgment is that the situation is typical, then typical options or standard operating procedures will be generated. But the judgment is that the case is

somewhat atypical, the decision maker will attempt to understand the situation by analyzing its features for a "best fit" to competing prototypes or schema.

We have identified four different types of information that are relevant to formulating a situation assessment: (1) plausible goals, (2) critical cues and causal factors, (3) expectancies, and (4) typical actions.

Plausible goals. Part of what it means to understand a situation is to understand what will be possible to accomplish. Goals here do not refer to the context-free goals of decision-theoretic models (generally the maximizing of some value), but to specific outcomes that a decision maker tries to achieve (i.e. the context-bound goals of Keeney & Raiffa, 1976).

Because the concept of "goals" can be vague, we suggest that a useful way to conceptualize goals is through contrast sets (Olson, 1970), since the selection of a goal also has implications about goals that were not selected. For example, an FGC could claim that the goal was to "do my job," but this is scarcely enlightening since there is no reasonable alternative. The more specific goal of "performing an interior attack" is meaningful because the alternatives are to "perform an exterior attack," "perform search-and-rescue," "call in a second alarm," or "abandon the effort and take precautions that the fire does not spread." In other words, the pragmatic meaning of a goal is communicated by contrasting it to alternative goals.

Critical cues and causal factors. In field settings, there are usually a great many events to attend to, and novices often feel overwhelmed with stimuli. Proficient decision makers do not feel this overload. The RPD model hypothesizes that the formulation of a situation assessment includes prioritizing critical cues, helping insure that attention is not diverted to less important cues or events.

9

For example, FGCs have learned to quickly scrutinize the color of flames. (This cue has meaning in terms of fire temperature and, by inference, the types of materials that are burning. Similarly, patterns of smoke convey information about the intensity of the fire by the pressure with which smoke is being pushed out of the building. Perceptual cues usually convey causal implications. Additional types of information also have causal implications. The nature of roof construction may imply vulnerability to damage, affecting the safety of sending firefighters onto the roof. However, there are times when the nature of roof construction is less important (e.g., when the danger is from smoke, not flames). The situation assessment includes attention to relevant cues and types of information.

Critical cues may also determine the timing of actions. In some of the fire ground decisions we studied, the decision maker's expertise was in recognizing when to act ("I held off ventilating the roof until I could see that the fire was beginning to spread to the attic").

Expectancies. Expectancies function to prepare decision makers for action and to provide clues for testing whether the situation is correctly understood. The situation assessment includes expectancies of what is likely to happen, and when. These expectancies can include a sequence of events, or a time course for events.

For an inexperienced decision maker, expectancies are poorly formed, vague, and hard to test. In contrast, an experienced decision maker holds clear expectancies. If events do not fit these expectancies, the resulting discrepancy raises questions about whether the situation assessment was correct. Thus, in one case we studied, an experienced FGC directed a stream of water at the area believed to be the seat of a fire. If correct, he

10

expected the pattern and color of the smoke to change within 20-30 seconds. When he saw no changes after about 45 seconds he suspected the seat of the fire was elsewhere.

The recognition of feasible goals includes expectancies about having sufficient time for accomplishing the goals. Goals are often linked to a timetable, which may be revealed in messages indicating whether plans are on schedule, or ahead of schedule, or "behind the power curve." If time expectations are violated it will create a feeling of urgency, and possibly a shift in situation assessment. For example, the FGC in the case cited above recognized that there was no longer enough time to extinguish the fire in the apartment building before it became a danger to the occupants. Smoke was already showing on the fourth floor, indicated that fire had spread to that area. There was no longer opportunity to direct water down at the fire.

Typical actions. A familiar situation evokes a familiar set of actions. It may even be useful to postulate an "action queue" of typical reactions to a situation at hand.[2] Therefore, the concept of situation assessment includes the identification of typical responses.

It has been asserted that proficient decision makers generate options on the basis of typicality--what is usually done in such a situation. Other possibilities are that an option is generated on the basis of recency (what was done the last time this came up), availability, or other factors. The generation of options by proficient decision makers is not fool-proof, and certainly does not invariably lead to optimality. The advantage is that it

_____

[2]Robert Holt (personal communication) has helpfully suggested the concept of an action queue in this context.

effort.

The recognition of a situation includes recognizing actions as well as recognizing goals and cues. One thing we learned from interviewing proficient decision makers is that much of their attention was on their own reaction, primarily the set of orders to give to their troops. In other words, a fire might feel like a "search-and-rescue" situation more than a "single-family 2-story home with brick exterior." The emphasis is on a functional understanding of what to do, as well as a structural understanding of cues and relationships.

Serial Evaluation

Serial evaluation refers to the assessment of options one at a time until a satisfactory one is found. Serial evaluation is different from concurrent evaluation of options, whereby a set of options is generated and evaluated comparatively.

This incident, taken from our interview files, illustrates serial evaluation.

> The commander of an emergency rescue squad arrived at the scene where a young woman, either drunk or on drugs, had either jumped or fallen from a highway overpass. She probably was attempting suicide by trying to fall to the highway below. But she missed, falling instead on a support strut for a highway sign. She was lying face-down on this strut, semi-conscious. A hook and ladder truck pulled up, and was directed to the highway below, to block traffic. Two members of the squad climbed out, and one pinned her legs to the strut while the other pinned her arms. The decision was how to raise her to safety without endangering the crew. The head of the squad told us that he first considered using a Kingsley harness which is the standard rescue equipment, but in imagining its use he could see that it would not work. Since it is attached from the front, he couldn't see how to push the woman to a sitting position without risk to all three people. He imagined attaching the Kingsley harness from the back, but

saw it would put excessive strain on the woman's back and
so rejected that option. Next, he thought of using a Howd
strap, which ties onto a victim (in contrast to a Kingsley
harness which snaps on), but saw that it ran into the same
problems so he rejected it. The next option he considered
was using a ladder belt—a heavy belt that firefighters
wear when climbing up several stories on a ladder, so that
they have a snap to attach to the top rung of the ladder
in case they lose their balance. He imagined lifting her
up a few inches, slipping the ladder belt under her waist,
buckling it closed (only 1 buckle is involved) and tying a
rope to the snap attachment. This is the option that he
selected, and the rescue was made.

In this example, the decision maker considered three options (five if you
count attaching the Kingsley harness and Howd strap from the back), but at no
time contrasted the strengths and weaknesses of one option versus another.
Instead, each option was examined in turn until a workable one was identified.
The decision maker reported that the whole decision took less than a minute.

The work of Simon (1955) is probably the best known discussion of a
serial generation and evaluation strategy in the behavioral science
literature. Simon described the use of satisficing as a means of quickly and
efficiently finding an effective option. Satisficing is a process by which
choices are evaluated one at a time until a satisfactory one is found, and
then implemented. It differs from an optimization strategy in which virtually
all options have to be generated and evaluated in order to determine which is
best. Simon claimed that efficient business executives relied on satisficing.
Cyert and March (1963) have also noted the importance of searching for the
first option that works instead of trying to find the best option.

The RPD model extends the concept of satisficing in several ways. It
asserts that the first option selected from the "action queue" is the most
typical option, and therefore has a high likelihood of being effective.
Therefore the proficient decision maker begins with a promising option, making

13

satisficing a more powerful strategy than if options were being generated randomly. The RPD model also asserts that options are selected roughly in order of their typicality, although other influences are undoubtedly present, such as recency and availability. Typicality itself is certainly linked to the concept of representativeness (Tversky & Kahneman, 1974), which under some conditions can mislead subjects. There is a trade-off here. Proficient decision makers have gained the ability to be generally successful and to react quickly. They accept the risk of being sub-optimal, and of occasionally having to spend time evaluating an option that is unacceptable.

A serial evaluation strategy as posited by the RPD model continuously makes available to a decision maker a preferred course of action. If time pressure forces a response, decision makers are prepared. In contrast, a concurrent evaluation model would leave a decision maker unprepared for action during the time course of the analysis. Only when all the analyses were completed would it become clear which course of action to select.

Progressive Deepening

Progressive deepening is the process of imagining how an option will be carried out within a specific situational context. It is the attempt to anticipate each important step, to notice the most likely reaction(s) to that step, to find the best way(s) to handle that reaction. It is an important component of recognitional decision making. Progressive deepening enables a decision maker to forecast the adequacy of a course of action. Within behavioral decision theory, options are evaluated by comparing them to each other with regard to how well they satisfy a set of criteria. In contrast, the RPD model asserts that one action is evaluated at a time. This is done by

14

imagining how the action would be implemented within the specific setting.[3] It is like running an "instant pre-play" to see if anything might go wrong.

An example might be useful. The head of a rescue unit arrived at the scene of a car crash. The victim had smashed into a concrete post supporting an overpass, and was trapped unconscious inside his car. In inspecting the car to see if any doors would open (none would), the decision maker noted that all of the roof posts were severed. He wondered what would happen if his crew slid the roof off and lifted the victim out, rather than having to waste time prying open a door. He reported to us that he imagined the rescue. He imagined how the victim would be supported, lifted, and turned. He imagined how the victim's neck and back would be protected. He said that he ran this imagining through at least twice before ordering the rescue, which was successful.

One of the first descriptions of this strategy in the psychological literature was in the work of de Groot (1965/1978). He coined the term "progressive deepening" to describe how chess grandmasters follow out a line of play and make sure it does not lead to any blunders. De Groot studied chess players trying to pick the best move in a difficult position. In the 40 protocols he presents, the chess players considered anywhere from 2 to 11 options but almost never compared one option to another.

We have expanded on de Groot's (1965/1978) work by hypothesizing that within the context of recognitional decision making, progressive deepening can help the decision maker in a number of ways: find weaknesses in an option; find ways to repair these weaknesses and thereby improve the option; discover

---

[3]Appreciation is expressed to Alexander Levis for discussions about the importance of imagery for decision making.

new opportunities that arise through implementing an option; alert the decision maker to previously ignored dynamics of the situation, thereby helping to modify the situation assessment.

In some ways, the topic of progressive deepening will overlap with the concept of contingency planning. However, contingency planning sometimes refers to systematic examination of plans. If the contingency planner checks for possible errors and oversights by examining as many assumptions as possible within the time available, this can be a very tedious process that could bog down in an exponential explosion of different factors and possibilities. In contrast, contingency planning by progressive deepening enables a skilled performer to be alert to important flaws in a plan without having to examine everything, and without having to decide what to examine and what to ignore (which entails first examining everything).

Progressive deepening also affects situational understanding. As actions are imagined, new features of the situation may be found, so Figure 1 shows an arrow leading back from the box "Imagine Action" to "Recognize the Situation."

III. IMPLICATIONS

The RPD model developed under this contract offers several important contrasts to normative and prescriptive models. What does the concept of recognitional decision making have to say about prescriptive decision models? First, let us examine such a model.

A strong position on prescriptive decision making has been taken by Janis and Mann (1977) who recommended that decision makers should be generating and contrasting options whenever possible. They claimed that decision making is stressful, that people avoid it when possible, and that many times where concurrent evaluation between options is appropriate and necessary, it does

16

not happen. Janis and Mann did not intend their advice for situations where there was extreme time pressure, but it is instructive to examine their ideas nonetheless.

For Janis and Mann (1977), there are seven criteria to be used in judging whether decision-making procedures are of high quality. They define the "ideal" decision maker (p. 11) as one who should:

-- thoroughly canvass a wide range of alternative courses of action;

-- survey the full range of objectives to be fulfilled and the values implicated by the choice;

-- carefully weigh whatever he knows about the costs and risks of negative consequences, as well as the positive consequences, that could flow from each alternative;

-- intensively search for new information relevant to further evaluation of the alternatives;

-- correctly assimilate and take account of any new information to which he is exposed, even when the information or judgment does not support the course of action he initially prefers;

-- reexamine the positive and negative consequences of all known alternatives, including those originally regarded as unacceptable, before making a final choice;

-- make detailed provisions for implementing or executing the chosen course of action, with special attention to contingency plans that might be required if various known risks were to materialize.

Janis and Mann (1977) assert that "failure to meet any of these seven criteria when a person is making a fundamental decision (one with major consequences for attaining or failing to attain important valu) constitutes

a defect in the decision-making process" (p. 11). They are well aware of the problems of making decisions under moderate time stress, and cite research showing that time pressure leads to a failure to make effective use of relevant and available information. They label the condition of "hypervigilance" as one where high conflict exists, along with a belief that a satisfactory solution exists, but with an apparent lack of time to search and deliberate. Here, the decision maker is hypothesized to display indiscriminate "openness" to all information. Janis and Mann (1977) repeatedly complain that people terminate information searches before all relevant data are examined.

The framework Janis and Mann (1977) are using is fairly typical of decision research that has attempted to formulate techniques to improve decision quality. For most of this work, the perspective taken is one described above as concurrent evaluation, and as analytical decision making. The decision maker is viewed as "faced with alternatives", which can be specified as branches emanating from a single point in a search tree. It s also natural to speak of the decision maker "considering the consequences" of each alternative in terms of an analysis of future states (odds/ probabilities) weighed against alternative goals (preferences/utilities). Techniques such as Decision Analysis and Multi-Attribute Utility Analysis (MAUA) have been derived to help the decision maker work out the various consequences of options.

This perspective leads to the conclusion that humans are limited and biased decision makers (e.g. Kahneman, Slovic, & Tversky, 1982; Nisbett & Ross, 1980). It has been a logical step, therefore, to focus decision support on methods of debiasing judgments (Fischhoff, 1982; Kahneman & Tversky, 1979),

18

calibrating probability estimates (Kadane & Lichtenstein, 1982), instructing in optimal combinatorial methods (Zakay & Wooler, 1984), and the like.

The potential significance of this work is enormous. If it is possible to develop general methods to improve decision making, then these methods could be trained and they could be embedded within decision aids to provide a large improvement in decision quality.

Unfortunately, the payoffs have yet to be seen. Decision aids, to support the use of Decision Analysis and MAUA, do not seem to have been well accepted in operational settings. With a few exceptions, decision training has not been shown to be very effective, and under time pressure such training has not shown any benefit (Rouse, 1978; Howell, 1984; Zakay & Wooler, 1984).

One way to understand the prescriptive implications of a recognitional decision model is to examine some of the standard recommendations for improving decision making.

Should Proficient Decision Makers Generate as Many Options as Possible?

From the perspective of recognitional decision making, the answer is "No."

This recommendation is heard from both decision researchers (e.g., Gettys, 1983) and practitioners writing popular books and articles (e.g., Janis & Mann, 1977). In the time-pressured environments we studied, there simply was not enough time to follow such advice. It takes time to generate and evaluate options, and delays may be intolerable. In addition, the situation may shift during the analyses so that the whole process has to start over again. Even in the absence of time pressure we rarely observed proficient decision makers trying to generate large sets of options. In our

research it is the novices who seem intent on generating options. Experts can recognize what to do right away and do not need to search further. Therefore, advice to generate large option sets is telling people to act like novices. The time to develop large option sets is when a situation is encountered that is unfamiliar, or when there are disputes.

Should Proficient Decision Makers Try to Rely on Analytical Methods, Including Concurrent Evaluation of Options, Wherever Possible?

Again, the answer is "No."

A recognitional decision strategy is a valid use of expertise in decision tasks and is a strong alternative to a generation/comparison strategy. In a recognitional strategy the expertise of the decision maker comes out in the identification of the appropriate option to consider, rather than in the way evaluation dimensions are selected or weighted or options are ranked.

The danger is not just that, by requiring proficient decision makers to perform analytical decision making and concurrent evaluation, they will then be forced into performing sub-tasks that are time consuming and inefficient. The greater concern is that they will be unable to make effective use of their own expertise. The Decision Analysis and MAUA approaches may not leave much room for the recognitional skills of experienced personnel. Therefore the risk of using these approaches is that decision performance will become worse, not better. In addition, trainees may not have a chance to develop expertise if they learn to rely on the analytical methods rather than developing their own recognitional capabilities. Remember that novices lack the recognitional skills needed to effectively perform recognitional decision making. It should not be difficult to convince novices to rely on analytical decision aids,

necessary for recognitional decision making.

Do Human Operators Have Too Many Judgment Biases to be Entrusted with Decisions?

The answer to this question is also negative.

Research on such biases as availability, representativeness, poor use of probabilistic data, and so on, has created an impression that people are inherently flawed decision makers. The implication is that we should develop training programs to reduce biases, rely on decision aids, and train special decision consultants.

These biases have been demonstrated in settings where context has been carefully limited, tasks are well-defined, and experience level is usually low. In other words, the opportunity for effective recognitional decision making has been limited, and attempts to apply recognitional decision making lead to errors. How well do these data generalize to actual decision tasks? Christensen-Szalanski and Beach (1984) has argued that these classical decision biases are artifacts of laboratory methodology and of the analytical perspective; they showed that studies of novice decision makers usually found evidence for biases whereas studies of experts usually documented their strengths. Christensen-Szalanski (1986) has also shown that judgment biases, even if they exist outside the laboratory are of little importance in decision making since the proportion of actual decisions they affect will be quite small. That is, the biases primarily operate on very low frequency events (e.g., a clear bias in the way physicians diagnose pneumonia would lead to an average of only one missed diagnosis per year). And if the frequency of such events increases, so will experience with them, thereby diminishing the bias.

21

Furthermore, Christensen-Szalanski cites a series of studies showing that "biased" decision processes do not lead to much reduction in the quality of the decisions. For example, miscalibration bias is clear cut but has almost no effect on accuracy of forecasting.

Furthermore, some tendencies that show up as biases for well-defined laboratory tasks may be of value in the field. "Biases" such as availability and representativeness reveal the fact that proficient decision makers have learned to rely on episodic memory. They can store earlier experiences as potential analogues to guide future performance. Surely the skilled use of episodic memory would be a strength for proficient cognitive performance in general, rather than a weakness for handling abstract story problems about female bank tellers (e.g., Tversky & Kahneman, 1983).

The optimization model has gained prominence as a normative model in psychology, despite its bases in statistical analysis and economic theory. Decision Analysis and Multi-Attribute Utility Theory with their emphasis on generating many options and then evaluating them systematically, has been likened to a process of evolution (Cooper, 1987). Success is achieved by starting with diversity and then applying stringent criteria for continuation. In studies with naive subjects, it has made good sense to encourage them to be creative and generate many options. Furthermore, if we have more faith in our conscious abilities to analyze and evaluate options than in the non-conscious abilities to generate the options, then we have a procedure where the important part is done in a way that can be controlled, observed, and improved. However proficient decision makers can rely on recognition to generate an option that is usually workable. We may not know how they do

this, and we may never be able to bring this skill under conscious control, but we should not ignore or deny it. And we should try not to develop prescriptions that interfere with recognitional decision making.

Summary

How can we characterize the major points of difference between recognitional and analytic decision making? There are seven important differences.

One difference is the focus on situation assessment rather than on the selection of an option. A recognitional model presupposes that the task has been recognized as being familiar in some important ways, so that the decision maker understands the plausible goals, cues and variables to monitor, expectancies, and the typical reactions. Descriptive models of analytical decision making have generally not addressed situation assessment. Prescriptive models like Decision Analysis regard situation assessment as a construction of states of the world, but have more to say about the consequences of selecting individual options.

A second difference is the mechanism for generating options. For analytical decision making, the assumption is of a fairly random process requiring careful evaluation. In contrast the RPD model describes how proficient decision makers can generate promising options as the first ones considered.

A third difference is that the RPD model concentrates on satisficing whereas analytical models have emphasized optimizing.

A fourth difference is about the nature of the evaluation. Analytical models deal with strong techniques for performing concurrent evaluation. The

23

RPD model describes decision makers as relying on progressive deepening to perform serial evaluation. The decision maker's task is seen as anticipating the outcome of implementing the typical reaction.

A fifth difference is in the treatment of options. The RPD model views options as being elaborated during the progressive deepening process. Limitations are found but the decision maker often tries to find ways of overcoming them, thereby strengthening the option. In contrast analytical models treat options as completed; attempts to modify and improve options would disrupt the evaluation process.

Sixth is the use of imagination. The evaluation process relies on the decision maker's ability to imagine how the option will be carried out, using world knowledge to anticipate pitfalls.

Seventh is response availability. Decision makers almost always will have an option that is ready to implement if time runs short; they only have to curtail their evaluation. In contrast an analytical strategy prevents a decision maker from knowing which option is favored until all the computations are completed.

What type of model is the RPD model? It is a descriptive model, derived from observations made in field studies. We primarily examined proficient decision makers, often driven by time pressures, but some of the studies examined less proficient decision makers and incidents that were not so time pressured. We have studied non-routine incidents, and in some studies we have probed only decision points that were non-routine, since we expected that such decisions would be most likely to require analysis. Even so, we found that the proficient decision makers were relying on recognitional strategies.

24

The RPD model is also a conceptual model, a framework for understanding how people function under operational conditions. In its present form it lacks the set of clear postulates that would allow it to be used to generate testable hypotheses. Thus, it is hard to imagine how the RPD model could be rejected. In the presentation of research findings, data were cited that supported the model but there was no presentation of negative research, because it is not clear what would constitute negative findings. For this reason, it anticipated that the model will have greater value for applied questions than for generating basic research.

Is the RPD model pertinent to decision making? Since the model describes processes where concurrent evaluation of options is avoided, and since it has bee claimed that they are the core of making a decision, the model may more appropriately be considered a description of problem solving than of decision making.

Berkeley and Humphreys (1982) treat decision making as concurrent evaluation. They define decision making as "the moment of choice among alternative immediate acts the decision maker has under consideration." (p. 203) This general type of definition is fairly standard in the field. Under this definition, we have not been studying decision making.

However, this definition may be too restrictive. It defines a phenomenon that is rarely encountered outside of laboratory conditions. The people we studied included fireground commanders with over 20 years of experience, and they claimed that they hardly ever used concurrent evaluation of options. Yet they were handling tasks that called for the allocation of personnel and equipment. And they were able to identify a number of "decision points" where reasonable options existed, options that someone with less experience might

25

well have chosen. Isenberg (1984) has also reported that business executives could recall few instances in their entire careers where they made decisions using concurrent evaluation about options. If these people are not making decisions, how relevant is the concept of decision making to applied psychology?

Of course, our data consist of verbal reports. It is quite conceivable that the people we studied were actually performing concurrent evaluation, but were doing it unconsciously or had forgotten about it. But the burden of proof shifts to the proponent of unconscious concurrent evaluation, to demonstrate that this phenomenon occurs and to explain how it is done. Until such proof is offered, there is no compelling reason to believe in the phenomenon of unconscious concurrent evaluation.

A more useful definition of decision making may be: identifying a course of action at a point where meaningful options exist. Under this definition, the decision maker does not have to consider more than one option actively. What makes it a decision is that meaningful options do exist and that the decision maker can articulate them if necessary. The focus here is on the task, not inside the head of the decision maker. It would allow us to compare how experts and novices perform the same task, contrasting their strategies. We could also study how changing the task conditions affected experts and novices differently. With the Berkeley and Humphreys (1982) definition it is natural for researchers to study naive subjects to make sure they are not using recognitional capabilities to avoid concurrent evaluation of options.

Lipshitz (1987) takes an even more extreme position than ours. He argues that decisions are fictions, artificially created under laboratory conditions. In naturalistic settings people function in a seamless web of intentions and

26

activities. He feels that it is a mistake to claim that individual decisions are being made, and that distinct goals are being identified. These concepts may be helpful in communicating, but they are also misleading. Looking at the flow of activities during a critical incident, it is clear that there are an infinite number of potential choices, and countless possibilities for decision points. Similarly, goals are not simply presented or deactivated. The actual situation may include a variety of intentions which will increase or decrease in importance, sometimes gradually and sometimes suddenly; it is misleading to segment out one or two intentions at the point where they were suddenly given prominence and pretend that these were the decision maker's goals.

Lipshitz's (1987) arguments are consistent with the observations we gathered during our interviews. We have adopted a less stringent definition than Berkeley and Humphreys (1982) because we judge their definition as too limited to be valuable for applied research. Lipshitz's position is even less stringent than ours but we will have to see how useful it is for providing direction for improving task performance. This is a key criterion--how helpful is each theoretical position to professionals working on applications? Theoretical hairsplitting can go on forever; guidance is needed today.

IV. REVIEW OF STUDIES

Our primary means of data collection was a Critical Decision method (Klein, Calderwood, & MacGregor, in press) developed under this and related contracts. It will be helpful in interpreting the specific results reviewed in this section to briefly outline key features of the method.

Overview of Critical Decision Method

The Critical Decision method is a retrospective interviewing strategy that shares many features with other methods, especially those related to

Flanagan's (1954) critical incident technique. Specific features of the technique include its focus on non-routine specific cases, and the use of probes to elicit information that may not be offered spontaneously.

The method was designed to strike a balance between a host of research objectives and practical constraints. For example, direct observations of command decisions coupled with an on-going verbal protocol of a commander's thought processes was first considered (see Hoffman's (1987) Method of Tough Cases). However, such an approach was deemed impractical in this case. Not only are challenging incidents relatively rare in any single location and expensive to cover because of the extreme time pressure, but the nature of the task makes any risk of outside interference untenable. We have used on-site observations to develop requisite domain knowledge prior to performing the actual elicitation task, and whenever possible to augment the data gathering.

At another extreme, simply asking fire ground commanders for unstructured accounts of their decisions would have resulted in little more than unrelated "war stories." Our goal was to focus the expert on those elements of an incident that most affected decision making, and to structure responses in a way that could be summarized along a specified set of dimensions while still allowing the details to emerge with the commander's own perspective and emphasis intact.

Core Procedures

The procedures adopted for the Critical Decision interviews represent our solutions to meeting these goals and practical considerations. The basic procedure can be summarized in the following steps:

Step 1: Select incident. Incidents were typically self-selected by the commander with the criterion that the case should represent a "command

28

challenge;" that is, they should illustrate a situation in which a decision had a significant impact on the outcome (either successfully or unsuccessfully). This selection criterion is common to critical incident methods (e.g. Klemp & McClelland, 1986) as a means of obtaining the most detailed and accurate reconstructions. In some cases incidents were selected by the interviewers or by supervisors during on-going operations.

Step 2: Obtain unstructured incident account. The participant is asked to describe the incident from its onset (e.g., the time he received the alarm) to the time when the incident was judged to be under control. For the most part this account proceeded without interruption by the interviewers, except for minor points of clarification. The procedure accomplished several goals. First, it created a context for understanding on the part of the interviewer. Second, the account served to activate the officer's memory of the event as a context for questioning. In addition, we judged that the procedure helped us achieve a high level of cooperation from the officers by establishing us as listeners rather than interrogators. During on-site observations this step might be very brief or eliminated.

Step 3: Construct incident timeline. After the incident had been related, the interviewer proceeded to reconstruct the account in the form of a timeline that established the sequence and duration of each event. Events included both objectively verifiable occurrences (e.g. "the second alarm equipment arrived two minutes later") and thoughts and perceptions reported by the officer (e.g. "the color of the smoke indicated the presence of a toxic substance," "I thought I might have to call a second alarm at this point"). The timeline served to establish a shared awareness of the "facts of the case." Many times inconsistencies in the account could be detected and

29

corrected on the basis of the timeline, and missing facts filled in. In
addition, questions about the timeline focused the officer's attention on
events from more than a single time perspective, an approach having
demonstrated utility for obtaining accurate eyewitness testimony (Geiselman,
Fisher, MacKinnon & Holland, 1985).

Step 4: Decision point identification. During the timeline
construction, specific decisions were identified for further probing. In some
cases the verbal cues marking a decision were obvious (e.g. "I had to decide
whether it was safe enough to send my crews inside"), but this was not always
the case. In other cases, it would be clear that an officer was taking one of
several possible courses of action or was making a judgment that affected the
outcome, but there was no clear indication that the officer saw himself as
"making a decision" at this point. A decision point was probed if the
participant confirmed that other reasonable courses of action were possible or
that another participant (perhaps one with less or greater expertise) might
have chosen differently.

Step 5: Decision point probing. Different studies have used different
probes, depending on the objectives of the projects. Interview Guides
included in the complete study reports indicate the wording of questions that
were systematically asked as part of the interview. Table 1 summarizes the
probe types that have been routinely used.

Questions to elicit the details of cue usage were almost always asked
first as part of the timeline construction, and represented the current
information that was likely to have been heeded at each event time. Prior
knowledge was also probed. We had a special interest in eliciting any recall

30

of prior experiences that influenced the officer's size-up or expectancies about a situation. Such specific remindings were coded as analogues.

Table 1

Critical Decision Interview Probes

Probe Type	Probe Content
CUES	What were you seeing, hearing, smelling...?
KNOWLEDGE	What information did you use in making this decision, and how was it obtained?
ANALOGUES	Were you reminded of any previous experience?
GOALS	What were your specific goals at this time?
OPTIONS	What other courses of action were considered, or were available to you?
BASIS	How was this option selected/other options rejected? What rule was being followed?
EXPERIENCE	What specific training or experience was necessary or helpful in making this decision?
AIDING	If the decision was not the best, what training, knowledge, or information could have helped?
TIME PRESSURE	How much time pressure was involved in making this decision? (Scales varied.)
SITUATION ASSESSMENT	Imagine that you were asked to describe the situation to a relief officer at this point, how would you summarize the situation?
HYPOTHETICALS	If a key feature of the situation had been different, what difference would it have made in your decision?

Study 1: A Critical Decision Study of Expert and Novice Fire Ground Command
Decisions

Goals. The goals of this study were to model the decisions made by experienced urban fire ground commanders (FGCs) using the Critical Decision method and to develop methods of analysis that would aid in understanding the role of experience in decision making.

Method. Critical decision interviews were carried out with 12 "Expert" and 12 "Novice" FGCs employed by six professional midwestern urban fire departments. Experts had an average of 19.5 years of fire fighting experience with an average of 11 years as an officer. Novices had an average of 10 years of fire fighting experience and less then two years as an officer.

Coding. Each decision was classified into one of nine types. The types were defined in terms of the intersection of two conceptual dimensions. Dimension #1 reflected serial or concurrent evaluation. This continuum is anchored on one end by choice involving little or no deliberation by the FGC. For these events, the FGC's actions appeared to be based primarily on his previous experience with similar events. When conscious deliberation did occur it frequently involved identifying and clarifying the nature of the situation itself or the specifics of action implementation or timing. These processes are commonly relegated to "predecision" stages or studied as aspects of monitoring or supervisory control, but we found them to be inseparable from decision making in this natural context. At the other end of the continuum were decisions fitting the definition of decision making more closely, in which action choices were deliberated in an attempt to meet multiple and sometimes conflicting goals. Dimension #2 reflected the degree to which the

(option) components of the decision problem.

Additional coding categories assessed whether a specific analogue had entered into the decision process, whether imagery was used, how quickly the decision was made, and whether the decision involved future planning.

Results. Fifty-four percent of the decision points were SA decisions. In these cases, identification and recognition of the situation allowed a choice of action to be generated and implemented without further deliberation. In 14% of the decision points, implementation and timing of a highly preferred or standard option was the most crucial issue. Even in the 32% of the cases that involved evaluation between options, 14% were serially evaluated. Thus, only 18% of the decisions fit the classical definition of decision making as concurrent evaluation between options.

Experts and Novices were roughly equally likely to deliberate about options. However, Experts used an approximately equivalent mix of serial and concurrent strategies whereas Novices appeared to rely more on concurrent deliberation. Experts were also more likely to deliberate about situational aspects of the decision problem, whereas Novices deliberated more about option implementation and timing. Experts also appear to construct novel option solutions much more frequently than Novices, and to report the use of imagery and evaluate potential options more frequently than Novices. Finally, Experts were almost twice as likely as Novices to consider future contingencies in their decision making.

Publication: Calderwood, R., Crandall, B., & Klein, G. A. (1987). Expert and novice fire ground command decisions (KATR-858(D)-87-02F). Yellow Springs, OH: Klein Associates Inc. Prepared under contract MDA903-85-C-0327) for U.S. Army Research Institute, Alexandria, VA.

Study 2: A Critical Decision Study of Distributed Decision Making in Wildland Firefighting

Goal. The goal of this study was to investigate decision strategies used by highly experienced commanders as they coordinated the efforts of thousands of firefighters during a large wildland fire. In this way we hoped to learn about decision-making strategies employed by command level experts in a high risk, often rapidly changing, distributed decision environment.

Coding. Coding followed the scheme developed in Calderwood, Crandall, and Klein (1987), except that the possibility that multiple decision strategies might be used was recognized by allowing multiple codes for a single decision point.

Method. This was an observational study carried out over eight days. Highly expert, command level, wildland firefighters working within the Incident Command System were observed and interviewed as they managed the suppression of a large forest fire. Seventeen very experienced members of two national Overhead Teams served as participants in this study. Critical Decision interviews were conducted by two on-site observers to determine the nature of the decision making strategies these experts used while performing their command-and-control activities.

Findings. As predicted, these experts relied heavily upon recognitional decision-making strategies. This was more pronounced in areas in which they had the greatest expertise. At many decision points they did not need nor have the luxury to deliberate among options. However, for decisions involving organizational issues and interpersonal negotiations (28% of the incidents identified as critical), we found a predominance of analytical strategies in which several options were evaluated concurrently.

Many of the complications of distributed decision tasks we had anticipated did not occur. There was little problem with information overload. Communication channels were limited but were used effectively. There was open communication about differences in the way situations were perceived and goals were formulated, but these were controlled so as to maintain team cooperation and morale.

Publication: Taynor, J., Klein, G. A., & Thordsen, M. (1987). Distributed decision making in wildland firefighting (KATR-858(A)-04F). Yellow Springs, OH: Klein Associates Inc. Prepared under contract MDA903-85-C-0327 for U.S. Army Research Institute, Alexandria, VA.

Study 3: A Critical Decision Study of Decision Making in Armored Platoon Command

Goals. This study was conducted to investigate the validity of laboratory based decision models for describing how Novices attain experience in armored platoon command. A primary focus was the Novice decision makers' description of contextual cues present at the time of the decisions. Investigators constructed a representational system for the cues and topics related by the Novice decision makers. These were compared to reports of the same decision situations related by experienced instructors who had observed evaluated them.

Method. CDM interviews were conducted with three classes of Armor Officer Basics over days three to six of field training exercises at Fort Knox, KY. Two observer/interviewers identified decision situations and interviewed student platoon leaders within twenty minutes of their completing the exercise scenarios. One observer rode in the platoon leader's tank during the exercises and then interviewed the trainee. The other interviewed the instructor who rode on top of the platoon leader's tank.

Coding. Three types of data were collected from the students in the interviews: the type of decision situation and decision strategy used, the cues and knowledge available to the student during the time of the decision, and self-performance ratings on a) tank and b) platoon actions as a result of the decisions. Instructors also reported cues and knowledge available and rated students' actions on the same performance scale presented to the students.

Results. The contextual cues and areas of knowledge students reported in their decision accounts were very similar to information offered by the

instructors. This suggests that performance difficulties were not generally the result of inattention to appropriate environmental cues but misinterpretation of the cues' importance. Analysis of the students' decision strategies revealed two main methods of resolution: 1) limited option deliberation and 2) recognition-primed decision implementation. These two methods were approximately equal in frequency. The students' high use of the latter type of strategy is consistent with our earlier research on more experienced personnel in other domains and supports the validity of a recognitional model for decision making at lower levels of expertise as well.

The number of analogues reported by the students was fairly stable across the observed training period and demonstrates that novices also use previous experience to guide decision making. Interestingly, analogues were helpful only about half of the time. On the remaining occasions the impact of analogues was mixed, ranging from neutral to disruptive.

One area in which some very interesting results surfaced was in the differential use of "hypotheticals" by the armored officer basic students (AOBs) used as compared to the more experienced track command instructors (TCIs). "Hypotheticals" reflected an evaluation of possible alternative future states. Overall, the AOBs showed a much weaker inclination to consider these hypotheticals. In addition, the more abstract the hypotheticals were, the greater the discrepancy between the number considered by the TCIs versus the AOBs. Terrain and factors concerning one's own tank were considered to be concrete hypotheticals, while platoon control, other friendly support, communications, and enemy unit hypotheticals were considered more abstract.

As mentioned earlier, the frequency of situational assessment statements was roughly equivalent for AOBs and TCIs. However, the primary area where a

discrepancy in SA did occur was in use of hypotheticals. While it does not appear that the AOBs are less attentive to SA information, it does seem that they are not yet able to select the most effective information to use to generate available options.

Publication: Brezovic, C. P., Klein, G. A., & Thordsen, M. (1987). Decision making in armored platoon command (KATR-858(B)-87-05F). Yellow Springs, OH: Klein Associates Inc. Prepared under contract MDA903-85-C-0327 for U.S. Army Research Institute, Alexandria, VA.

Study 4: How Do People Use Analogues to Make Decisions

Goal. The purpose of this project was to examine the data we had gathered using Critical Decision method to learn more about analogical reasoning and its role in decision-making.

Method. Data from Critical Decision interviews collected in five studies was reexamined for evidence of analogue use. The studies included two studies of urban fire ground command (Calderwood et al. 1987; Klein, Calderwood, & Clinton-Cirocco, 1986), a study of decision training during tank platoon exercises (Brezovic, Klein, & Thordsen, 1987), a study of decisions made during a wildfire incident (Taynor, Klein, & Thordsen, 1987) and a study of Air Force design engineers (Klein & Brezovic, 1986).

This data base contains over 400 decision points in all, culled from interviews with over 100 decision makers. The data were analyzed and compared in order to gain a broader understanding of the role of analogical reasoning in decision making. A total of 33 analogues were identified in enough detail to analyze the functions served by the analogue.

Results. Three functions of analogical reasoning were identified, 1) understanding situational dynamics, 2) generating options, and 3) evaluating the probable success or failure of implementing an option. Several tentative conclusions were also offered:

* Analogical reasoning is reported relatively infrequently by experts, perhaps because the individual cases have often merged into prototypes.

* When analogical reasoning occurs, it is often critical for success. For experts, it often emerges during non-routine cases.

* Novices appear to rely more heavily than experts on analogical reasoning, but have not learned how to apply the analogues, modify them, or

reject them. Therefore almost half the analogue use by novices results in poor choices.

Publication: Klein, G. A., & Calderwood, R. (1988). How do people use analogues to make decisions? Case-Based Reasoning Workshop, sponsored by Defense Advanced Research Projects Agency (DARPA), Clearwater Beach, FL.

Study 5: The Effect of Time Pressure on Expert Decision Making

Goal. The study investigated the effect of time pressure on the
decisions made by chess players at two different levels of skill. The
hypothesized results were based on assumed differences in the temporal
requirements of calculational and recognitional modes of processing.
Calculational processes, such as generating move-countermove sequences in
order to evaluate outcomes, are relatively time dependent. When time
constraints are imposed, calculations must be either truncated or omitted,
thereby impairing performance. Recognitional processes, on the other hand,
are defined as rapid and holistic. Performance based on recognitional
processes should therefore be relatively insensitive to time constraints.
Thus, we anticipated an interaction between time pressure and playing skill on
move quality in chess. An obtained interaction would provide converging
evidence for the claim the highly-skilled decision makers rely more on their
recognitional abilities than do less skilled individuals.

Method. The rated quality of moves for very strong (masters) and weaker
(class B) players was compared for tournament games played under regulation
(at least 50 moves in two hours) and blitz (6 minutes total playing time per
player) time rules. Tournaments were arranged as double round-robins wherein
each of three players at each skill level played each of the other players
four times, twice for regulation and twice for blitz games. This design
resulted in 24 games, six in each of the conditions resulting from crossing
the player-skill and game-type factors. Moves were rated for quality on a
5-point scale by a chess grandmaster.

Results. Results of the analysis of move quality ratings supported the
predicted interaction between skill level and game type. That is, the

decrement in move quality for blitz games compared to regulation games was greater for the class B players than for the masters. The validity of the interaction was supported by the fact that masters were more able to maintain higher quality moves in the blitz condition at the same time that they generated a substantially greater number of moves, and proportionately more complex moves, than the class B players. These results were interpreted as supporting the view that more highly skilled players can rely more extensively on rapid recognitional processes than less-skilled players. Of course, this does not rule out the possibility that given adequate time, more highly skilled players may also calculate more extensively and more profitably than weaker players.

Publication: Calderwood, R., Klein, G. A., & Crandall, B. W. (In press). Time pressure, skill, and move quality in chess. American Journal of Psychology.

Study 6: Classification of Decision Events

Goal. Over 50 published articles were reviewed relevant to event classification, prototypes, schemas, skill development, and "pre-decisional" decision processes. The goal was to synthesize the findings in order to suggest a framework for studying situational assessment processes.

Findings. Several surprising "holes" in our knowledge relevant to these areas were uncovered and summarized.

* There are surprising few links between the decision making literature and the literature on concept formation and categorization. Yet, the RPD framework suggests that a major component of decision making is in how an event is understood and classified.

* Little is known about natural event classification. Classification stimuli have tended to be objects or unidimensional concepts.

* The closest analogue to event classification may be in problem-solving studies which have used psychological scaling techniques to uncover and represent the "dimensions" on which similarity judgments are made. Many of these studies have compared the derived representations of Expert and Novice performers in order to draw inferences about the nature of skill development. However, these investigations have not considered how context may influence the judgments on which the clustering solutions are based. Nor have they considered how context might interact with skill.

Publication: Based on this literature review, a dissertation was proposed by the second author and accepted by the Psychology Department of the University of New Mexico. The study results will be sumbitted for publication.

43

Study 7: Protocol Analysis of Expert/Novice Fire Ground Decision Making During Simulated Incidents

Goal. The goal of this study was to experimentally demonstrate the suggestive findings obtained in previous studies in this series. Whereas the previous investigations have relied on retrospective interviews to probe for information, the present study obtained think-aloud protocols during simulated incidents. Thus, this approach would offer the first opportunity to judge the content and strategy differences of Expert and Novice decision makers unconfounded by differences in the type of incident and the information available in the situation.

The study was designed to address several inter-related issues of relevance to RPD model development:

* Does this alternate method provide convergent evidence for the serial evaluation strategy described by the RPD model?

* Does the method provide a technique for examining progressive deepening and imagery as a means of option evaluation?

* What aspects of situation assessment are spontaneously reported -- what cues, inferences, and goals are associated with command decisions?

* To what extent are these factors associated with domain expertise?

Method. Three simulations of fire ground incidents were developed. Key events of actual incidents were recreated using an audio-visual format to present the details of the incident from the commander's perspective. The simulation presents relevant radio communication and a series of graphic slides of an incident from the time of the initial alarm to a point where the incident has been brought under control. All events are depicted from the point of view of the FGC. A narrator supplies needed background information

44

that would be known to the commander or would become available in other ways during an actual incident. Key events are portrayed in near real-time. In their final form, each simulation contains multiple decision points that span the duration of the incident. The tape is stopped at these points allowing the participant, assuming the commander's role, to think-aloud about any decisions he would make at this point.

Twenty-two professional firefighters participated, 11 Experts and 11 Novices. Expert/Novice ranking was made on the basis of overall command experience and an ability rating made by the Chief of Suppression Officer of a major urban fire department from which participants were drawn.

Coding. A coding procedure was developed and tested for inter-coder reliability. The method classified each protocol remark into 12 independent categories related to cues, knowledge, actions, and goals. In addition, evidence for both RPD and concurrent decision strategies, imagery and analogue use, progressive deepening, and possible errors were noted when they occurred.

Results. Analysis of the frequencies of the remark categories substantiated the hypothesized differences in decision "focus" for Experts and Novices. Experts appeared to pay more attention to assessing the situation (noticing cues and making inferences based on the cues), whereas Novices pay relatively more attention to generating and evaluating options.

A content analysis based on a conceptual node graph of the remark categories was performed. This graphing method proved to be a powerful interpretive tool for abstracting within group commonalities and highlighting between group differences. The node graphs supported the idea that both situation assessment and action schemas were richer and more elaborated in the

45

Expert group. It also revealed underlying "decision points" that were difficult to detect in individual protocols.

Publication: The full report (Protocols Analysis of Expert/Novice Command Decision Making During Simulated Fire Ground Incidents by R. Calderwood, B. W. Crandall, & T. H. Baynes) of the study is included as Appendix I of this final report.

REFERENCES

Barlett, F. C. (1932). Remembering. Cambridge, England: Cambridge University Press.

berkeley, D., & Humphreys, P. (1982). Structuring decision problems and the 'bias heuristic'. Acta Psychologica, 50, 201-252.

Brehmer, B., & Allard, R. (1986). Dynamic decision making: A paradigm and some experimental results. Sweden: Uppsala University.

Brezovic, C. P., Klein, G. A., & Thordsen, M. (1988). Decision making in armored platoon command (KATR-858(B)-87-05F). Yellow Springs, OH: Klein Associates Inc. Prepared under contract MDA903-85-C-0327 for U.S. Army Research Institute, Alexandria, VA.

Calderwood, R., Crandall, B., & Klein, G. A. (1988). Expert and novice fire ground command decisions (KATR-858(D)-87-02F). Yellow Springs, OH: Klein Associates Inc. Prepared under contract MDA903-85-C-0327) for U.S. Army Research Institute, Alexandria, VA.

Christensen-Szalanski, J. J. J. (1986). Improving the practical utility of judgment research. In B. Brehmer, H. Jungerman, P. Lourens, and G. Sevon (Eds.), New directions in research on decision making. North Holland: Elsevier.

Christensen-Szalanski, J. J. J., & Beach, L. R. (1984). The citation bias: Fad and fashion in the judgment and decision literature. American Psychologist, 39, 75-78.

Cooper, W. S. (1987). Decision theory as a branch of evolutionary theory: A biological derivation of the savage axions. Psychological Review, 94, 395-411.

Cyert, R. M. & March, J. G. (1963). A behavioral theory of the firm. Englewood Cliffs, NJ: Prentice-Hall.

de Groot, A. D. (1965/1978). Thought and choice in chess (2nd ed.). NY: Mouton.

Edwards, W. (1962). Dynamic decision theory and probabilistic information processing. Human Factors, 4, 59-74.

Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. A. Slovic, & A. Tversky (Eds.), Judgment Under Uncertainty: Heuristics and Biases. New York: Cambridge University Press.

Flanagan, J. C. (1954). The critical incident technique. Psychological Bulletin, 51, 327-358.

Geiselman, R. E., Fisher, R. P., MacKinnon, D. P. & Holland, H. L. (1985). Eyewitness memory enhancement in the police interview: Cognitive retrieval mnemonics versus hypnosis. Journal of Applied Psychology, 70, 2, 401-412.

Gettys, C. F. (1983). Research and theory on predecision processes (TR 11-30-83). Norman, OK: University of Oklahoma, Decision Processes Laboratory.

Hoffman, R. R. (1987). The problem of extracting the knowledge of experts from the perspective of experimental psychology. AI Magazine, 8, 53-67.

Howell, G. E. (1984). Task influence in the analytic intuitive approach to decision making final report (ONR Contract N00014-82 C-001 Work Unit (NR197-074)). Houston, TX: Rice University.

Isenberg, D. J. (1984). How senior managers think. Harvard Business Review, 6, pp. 80-90.

Janis, I. L., & Mann, L. (1977). Decision making: A psychological analysis of conflict, choice, and commitment. NY: The Free Press.

Kadane, J. B., & Lichtenstein, S. (1982). A subjectivist view of calibration (82-6). Eugene, OR: Decision Research.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.), (1982). Judgment under uncertainty: Heuristics and biases. Cambridge, MA: Cambridge University Press.

Kahneman, D., & Tversky, A. (1979). Intuitive predictions: Biases and corrective procedures. TIMS Studies in Management Science, 12, 313-327.

Keeney, R. L., & Raiffa, H. (1976). Decisions with multiple objectives: preferences and value tradeoffs. New York: Wiley.

Kintsch, W., & Van Dijk, T. A. (1975). Comments on summaries of stories. Langages, 40, 98-116 (in French).

Klein, G. A. (In press). Recognition-primed decision. In W. Rouse (Ed.), Advances in Man-Machine Systems Research, 5. Greenwich, CT: JAI Press, Inc.

Klein, G. A., & Brezovic, C. P. (1986). Design engineers and the design process: Decision strategies and human factors literature. Proceedings of the 30th Annual Human Factors Society, 2, 771-775. Dayton, OH: Human Factors Society.

Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (1986). Rapid decision making on the fire ground. Proceedings of the Human Factors Society 30th Annual Meeting, 1, 576-580. Dayton, OH: Human Factors Society.

Klein, G. A., Calderwood, R., & Clinton-Cirocco, A. (1988). Rapid decision making on the fire ground (KA-TR-84-41-7). Yellow Springs, OH: ARI Technical Report, 796, ADA 199-492.

Klein Associates Inc. Prepared under contract MDA903-85-G-0099 for the U.S. Army Research Institute, Alexandria, VA.

Klein, G. A., Calderwood, R., & MacGregor, D. (In press). Critical decision method for eliciting knowledge. IEEE Transactions, Man, Systems and Cybernetics.

Klein, G. A., & Weitzenfeld, J. (1982). The use of analogues in comparibility analysis. Applied Ergonomics, 13, 99-104.

Klemp, G. O., & McClelland, D. C. (1986). Executive competence: What characterizes intelligent functioning among senior managers? In. R. J. Sternberg & R. K. Wagner (Eds.), Practical intelligence: Nature and origins of competence in the everyday world. MA: Cambridge University Press.

Lipshitz, R. (1987). Decision making in the real world: Developing descriptions and prescriptions from decision Maker's retrospective accounts. Unpublished manuscript. Boston University, Center for Applied Sciences.

Nisbett, R., & Ross, L. (1980). Human inference: Strategies and shortcomings of social judgment. In B. Brehmer, H. Jungermann, P. Lourens, and G. Sevon (Eds.), New Directions in Research on Decision Making. The Netherlands: Elsevier Science Publishers B.V.

Olson, D. R. (1970). Language and thought: Aspects of a cognitive theory of semantics. Psychological Review, 77, pp. 257-273.

Rapoport, A., & Wallsten, T. S. (1972). Individual decision behavior. In P. H. Mussen and M. R. Rosenzweig (Eds.), Annual review of psychology. Palo Alto, CA: Annual Reviews Inc.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the

internal structure of categories. Journal of Cognitive Psychology, 7, 573-605.

Rouse, W. B. (1978). Human problem solving performance in fault diagnosis task. IEEE Transactions on Systems, Man, and Cybernetics, 4, SMC-8, 258-271.

Simon, H. A. (1955). A behavioral model of rational choice. Quarterly Journal of Economics, 69, 99-118.

Taynor, J., Klein, G. A., & Thordsen, M. (1988). Distributed decision making in wildland firefighting (KATR-858(A)-04F). Yellow Springs, OH: Klein Associates Inc. Prepared under contract MDA903-85-C-0327 for U.S. Army Research Institute, Alexandria, VA.

Tversky, A. (1977). Features of similarity. Psychology Review, 84, 327-352.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. Science, 185, 1124-1131.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. Psychological Review, 90(4), 293-315.

Weitzenfeld, J. (1984). Valid reasoning by analogy. Philosophy of Science, 51, 137-149.

Wohl, J. G. (1981). Force management decision requirements for Air Force tactical command and control, IEEE Transactions on Systems, Man, and Cybernetics, 618-639.

Wohl, J. G., Entin, E. E., Kleinman, D. L., & Pattipati, K. (1984). Human decision processes in military command and control. In W. B. Rouse (Ed.),

Advances in Man-Machine Systems Research, Vol. 1, (pp. 261-307).

Greenwich, CT: JAI.

Zakay, D., & Wooler, S. (1984). Time pressure, training, and decision

effectiveness. Ergonomics, 27, 273-284.